

DOCUMENT RESUME

ED 162 901

SE 025 430

AUTHOR Burstein, Leigh
 TITLE Implications from the Beginning Teacher Evaluation Study for the IEA Seccd Mathematics Study.
 PUB DATE 78
 NOTE 25p.; Paper presented at the annual meeting of the American Educational Research Association (Toronto, Canada, March 27-31, 1978); Contains occasional light and broken type
 EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
 DESCRIPTORS Achievement; *Classroom Observation Techniques; Classroom Research; *Educational Assessment; Elementary Secondary Education; *International Studies; *Mathematics Education; *Teacher Behavior

ABSTRACT

The International Education Association plans for the Second Mathematics Study are deemed inadequate for the identification of the effects of instruction on mathematics education. Suggested modifications are based on the results of two studies. The first is a discussion of major methodological limitations of the Six Subjects Survey. The second is a comprehensive classroom process measurement used by the Beginning Teacher Evaluation Study which included extensive classroom observation and teacher reports of allocated time for specific pupils in selected content areas. (MF)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED162901

AERA
MAY, 1978

IMPLICATIONS FROM THE BEGINNING TEACHER EVALUATION
STUDY FOR THE IEA SECOND MATHEMATICS STUDY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Leigh Burstein

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

Leigh Burstein

University of California, Los Angeles

Graduate School of Education
University of California, Los Angeles
405 Hilgard Ave.
Los Angeles
California 90024

SE 025 430

Implications from the Beginning Teacher Evaluation
Study for the IEA Second Mathematics Study

Leigh Burstein

Department of Education

University of California, Los Angeles

As IEA begins the Second Mathematics Study, it seems appropriate to consider what can be learned from the previous IEA studies and related work about viable methodology for identifying educational effects. I would like to take advantage of this opportunity for 100% hindsight and comment on what I believe to be the major methodological shortcomings of the school effects analyses in the Six Subject Survey with special emphasis on the Science (Comber and Keeves, 1973) and Technical (Peaker, 1975) Volumes. I do this in order to identify those areas in which current plans for the Second Mathematics Study (Travers, 1978) need modification and augmentation.

My primary concern is that the present IEA plans for the assessment of classroom processes are inadequate for the identification of the effects of instruction on mathematics achievement and cannot be improved substantially within the survey framework. Evidence of the value of micro-assessment of classroom processes from the Beginning Teacher Evaluation Study (BTES) conducted at Far West Laboratory (FWL) (Berliner, et al., 1976, see attached annotated bibliography of BTES Reports) is discussed. The BTES work, and particularly the FWL procedures, have implications for the assessment of classroom procedures in the IEA Second Mathematics Study. In the concluding section, I consider how the FWL procedures can be adapted for IEA use.

Major Methodological Limitations of the Six Subjects Survey

Without a doubt, the Six Subjects Survey and its reports (Carroll, 1975; Comber and Keeves, 1973; Lewis and Massad, 1975; Passow, Noah, and Eckstein, 1976; Peaker, 1975; Purves, 1973; Thorndike, 1973; Torney, Oppenheim and Farnen, 1975; Walker, 1976) were an enormous undertaking and an outstanding contribution to the literature on the effects of schooling. The studies have been used to support and contradict a broad variety of theories about curriculum planning, retentivity, the role of family background and so on (See, e.g., Purves and Levine, 1975; Postlethwaite et al. 1974).

Perhaps, the research topic on which the Six Subjects Survey has had the greatest impact is the study of the determinants of educational achievement using regression-based analyses of survey data (Burstein, 1976; Burstein and Smith, 1977; Coleman, 1975; Finn and Kotlewski, 1977; Inkeles, 1976) and the corollary educational production function literature (Jau, 1977). In the most frequently cited follow-up study of IEA data, Coleman (1975) frankly admits that the variation in effects across the IEA studies caused him to question the wisdom of the overemphasis on verbal achievement in his Equal Educational Opportunity Report (Coleman, et al., 1966) and to suggest that school effects may be stronger in other content areas. Coleman's admission alone provided sufficient justification for considering the Six Subjects Survey to have been successful.

Though the studies may stand the test of time, certain aspects of their methodology can be retrospectively viewed as questionable.

The major methodological shortcomings included (a) the failure to match students with their own teacher(s), (b) the lack of a proper surrogate for entering ability, (c) the quasi-longitudinal block-wise entry of variables in the regression analyses, (d) the focus on school-level analyses and (e) the inadequate assessment of teacher, classroom, and school characteristics and processes. As a whole the set of shortcomings tended to accentuate

the influence of what was labeled as home background conditions (Sex and Age were inappropriately contained in this "block" of variables) and underestimate the effects of school-related conditions. The biases associated with the aggregation of the data (Burstein, 1976; Burstein and Smith, 1977), with the attribution of commonality to background by the fixed order of entry (Schwille, 1973), with the inadequate measurement of teacher/classroom/school processes, and with the misspecification of the model by excluding relevant variables such as ability all work in the IEA studies to reinforce the (mis)perception that schooling has little or no influence on educational achievement especially when compared to family background.

The Second Mathematics Study Viewed from the Perspective of the Limitations of the Six Subjects Survey

If the Second Mathematics Study is carried out as planned (Travers, 1978), many of the methodological shortcomings of the Six Subjects Survey will not be repeated. In the first place, the International Mathematics Committee (IMC) has determined that the study will be an "enterprise of, by and for the learning of mathematics in schools" (Travers, 1978; p. 2). Furthermore, the IMC has decided to "focus upon the classroom, what it is that teachers do there, and what are the associated student outputs." (Travers, 1978; p. 6). This choice of emphasis has meant a shift away from the school as central focus ((d) above), a determination to match student data with their specific teachers (point (a)), the decision to include a pretest at least for Population A (point (b)) and to focus on the factors influencing achievement growth (point (C)). Though more could be done to ameliorate the effects of each of these shortcomings, the current plans represent a substantial improvement over the earlier studies.

The IMC has also taken some steps to improve the assessment of classroom processes. This is not surprising given their emphasis on the classroom.

Their intent is to collect data on what teachers say they do as they teach mathematics. In particular, the assessment of classroom processes will be done primarily with survey instruments completed by teachers on their instructional strategies and the content they cover (opportunity-to-learn). Current versions of these questionnaires are already far superior to previous IEA efforts in this area. There is also some indication of support for the inclusion, at least as international options, of a limited amount of classroom observation and a few case studies to clarify further the nature of classroom processes.

Unfortunately, I doubt that the questionnaires alone will succeed in assessing classroom processes for the following reasons. It seems to me that the major complications in efforts to identify the effects of education has been our inability to develop adequate theories of educational processes within groups (classrooms and schools) and adequate methodology for analyzing the effects of such processes (Burstein, 1977, 1978).

Learning is more than an academic exercise based solely on instructional method, time-on-task and pupil ability. Learning takes place in classrooms filled with students and managed by teachers. Thus learning is also a social process that occurs in a social context. Teachers have ideas about how to organize instruction, how children learn, and how they view their own role as a teacher and as an individual in an organization. Children have motives of their own -- friendships, peer recognition -- which may be incidental, or even antithetical, to academic learning. The point is that unless the purposes and practices conducive to academic performance are also in line with teacher's goals and strategies and the students' aspirations, the social dynamics of the classroom can overrule the academic methods.

If learning in classrooms is both an academic and a social activity, and if teachers and students are both active participants with potentially

different perspectives, goals, and motives, then questionnaires soliciting instructional strategy and opportunity-to-learn information strictly from teachers are doomed to be inadequate. The teacher cannot be sufficiently dispassionate nor sufficiently perceptive to represent his/her own motives and actions accurately or to capture or reconstruct the behavior and rationale of students. Classroom observations, interviews with teachers and students and ethnographic descriptions of classroom activities would all seem to be necessary adjuncts to current plans to do justice to the role of classroom processes in learning.

It is true that the types of assessment of classroom processes being advocated are much more difficult to incorporate in the IEA study than questionnaires. They are time-consuming and require more extensive training and more intrusion in the classroom. Moreover, the political problems associated with intensive examination of classroom practices must be acknowledged.

Nevertheless, if the IEA studies truly represent earnest efforts to understand the factors affecting educational achievement, then an attempt to improve the measurement of classroom processes seems warranted. Furthermore, the inclusion of the more comprehensive techniques of measuring classroom processes can lead to further cross-national development of the capabilities to examine the educative process, another long-standing IEA goal.

Selected data from the Far West Laboratory (FWL) Beginning Teacher Evaluation Study (BTES) are presented below in effort to highlight the potential value of more comprehensive classroom process measurement. The study included extensive classroom observation and teacher reports of allocated time for specific pupils in selected content areas. In the pilot phase of the work, ethnographic studies were conducted in several second and fifth grade classrooms. In what follows, we focus on observation and teacher log data in fifth grade classrooms.

FWL Beginning Teacher Evaluation Study, Phase IIIB

The Beginning Teacher Evaluation Study (BTES) is a six-year research project being conducted by the California Commission for Teacher Preparation and Licensing (CCTPL) through a contract from the National Institute of Education. Far West Laboratory for Educational Research and Development (FWL) is currently conducting the investigation for CCTPL which is now in Phase IIIB. The purpose of BTES is to determine those aspects of the teaching process which interact with student characteristics to affect educational performance.

In their two-year study, FWL developed and investigated a model of learning whereby a pupil's achievement is a function of pupil aptitude, motivation, and the amount of active learning time (ALT) s/he received. Active Learning Time is defined as the time a student is engaged in material relevant to the content to be learned that is at an appropriate level of difficulty.

In order to obtain reasonable information about ALT, FWL devised thorough and sensitive classroom process instrumentation. The key instruments were teacher logs which report, on a daily basis throughout the academic year, allocations of the instructional time in specific content areas for a sample (6) of their pupils, and a time-sampling observational instrument which observers (1 or 2 per occasion, 1 entire class day each week of the academic year) used to collect information on allocated and engaged time on specific instructional topics, difficulty of material, instructional settings, and learner and instructor moves. Observational data were collected for the same sample of children for whom the teachers kept logs. Over the school year (1976-77), there were approximately 22 days of observation per classroom.

The FWL study included approximately 25 second grade and 25 fifth grade classrooms. All students in the participating classrooms were given a selection test covering reading and mathematics and child demographic data were obtained.

The test results were used to select a sample of six children per classroom whose instruction would be followed throughout the year. Children were chosen at random from those in the classroom whose scores fell between the 10th and 90th percentile of children in all same-grade classrooms on the selection test. The sampled children were given an extensive battery of tests covering reading and math content objectives (26 reading and 14 math at second grade; 19 reading and 16 math at fifth grade) in the early fall (A testing), in December (B testing), in May (C testing) and again at the beginning of the next academic year (D testing). A series of affective instruments were also administered.

It is impossible to summarize adequately the 30 or so reports produced by FWL. The material is both wide-ranging and deep in scope. It is clear that there is a high degree of variability in instructional activities, active learning time and achievement across subject matter and grade both among and within classrooms. For example, the second-grade teacher log reports for the A-B period indicate a range of 20 to 810 minutes in time allocations to addition and subtraction without regrouping with the largest within-class variation being from 230 to 630 minutes. The corresponding data for the longer B-C period were 60 to 1790 minutes for all pupils and 580 to 1790 minutes within a single classroom. If these differences in time allocations are accurate reflections of ALT and ALT is crucial to learning, then there are either substantial differences in students' needs for instruction in basic arithmetic operations, substantial differences in students' learning as a result of instruction, or both.

Effects of ALT on Achievement. Substantial evidence for the influence of ALT on achievement is also presented. FWL (Marliave, Fisher & Dishaw, 1977) ran a series of regression analyses of the following form:

$$ACHB = f(ACHA, ADJALLOC)$$

where ACHB = score on test in a specific content area in a given grade at time B

ACHA = score on test at time A

ADJALLOC = teacher's report of total allocated time to specific student in specific content area during A-B period, adjusted for student engagement rate determined from observational data.

The results for fifth graders in multiplication, division and fractions, the three content areas with the largest amount of instructional time in this grade, are presented in Table 1. It can be seen that the single measure of time makes a substantial contribution to achievement even after controlling for what should be the most powerful source of background, prior achievement.¹ Moreover, the magnitude of the standard deviation suggests that the relationship as presented is understated since the adjusted time allocations are not normally distributed.

Table 1 about here

General Allocation Data. Other data about the general patterns across fifth grade classrooms during the B-C period are worth describing. The observational data (Fisher, Filby & Marliave, 1977) indicated that on the average, mathematics was taught for 44 minutes (standard deviation = 13) out of an average 361 minute school day. Average daily time across classes varied from 20 minutes to 73 minutes. On the average, students were judged to be engaged in mathematics instruction 74% (standard deviation = 13) of the time with engagement rates ranging from 50% to 87% of the time. Self-paced mathematics instruction time averaged 76% with students engaged 72% of that time. In contrast, the

Table 1. Effect of Engagement Adjusted Teacher Allocated Time on Achievement Controlling for Prior Achievement (A-B Period, Grade 5, N = 122)

Content Area	Time in Content Area		Raw Regression Coefficient	t Statistic	Contribution to Total R ² After Preachievement
	Mean	Standard Deviation			
Multiplication	126.8	136.3	.0156	4.46	10%
Division	118.8	186.9	.0056	2.67	4%
Fractions	117.2	108.4	.0084	5.25	18%

Source: Marliave, Fisher & Dishaw, 1977

average engagement rate in other paced activities was 85%. About 60% of the instructional time was devoted to medium difficulty material with 35% easy material.

As Table 2 indicates, average daily time allocations, engagement rates, and difficulty of material varied across subject matter during the B-C period. The number of students receiving instruction in a specific area also varied.

Table 2 about here

Time Variation across Classrooms. The data over all classrooms masks class-to-class differences in time allocation, both teacher-reported and observed; engagement rates and time in specific content areas. Table 3 provides data from a sample of the 25 classrooms chosen for a variety of reasons: high or low observer daily time (classrooms 3, 4, 24, 25, 14), large variability of time within classrooms (classrooms 8, 17); high and low engagement rates (classrooms 23, 8 and 6), and large disparities between teacher's report and observers' reports of allocated time (classrooms 6 and 18). The classrooms have been ordered in terms of average daily time in mathematics reported by the observers.

Table 3 about here

Table 2. Number of Students, Average Daily Time, Engagement Rates, And Difficulty in Selected Content Areas in Fifth Grade Mathematics During the B-C Period.

Content Area	Number of Students Receiving Instruction	Average ^a Daily Time	Engagement ^a Rate	% ^a Easy	% ^a Medium
Fractions	99	11 (10)	73 (19)	24 (26)	72 (27)
Division	115	9 (8)	78 (15)	42 (36)	54 (34)
Multiplication	108	6 (6)	75 (24)	47 (35)	51 (35)

Source: Fisher, Filby & Marliave, 1977.

^a Standard deviations in parentheses.

Table 3. Observer and Teacher Reported Allocated Time, Engagement Rates, Proportion of Time Allocated to Specific Content Areas in Selected Fifth-Grade Classrooms During the B-C Period.^a

Class	Observer ^b Daily Allocated Time	Teacher ^b Daily Allocated Time	Engagement ^b Rate	Percentage of Time on		
				Fractions	Division	Multiplication
3	<u>20</u> (3)	22 (1)	83 (3)	5	20	22
4	<u>29</u> (2)	29 (1)	84 (7)	18	--	<u>49</u>
8	<u>33</u> (<u>13</u>)	24 (5)	<u>87</u> (6)	25	13	25
18	<u>43</u> (4)	<u>56</u> (--)	67 (<u>12</u>)	<u>52</u>	<u>33</u>	5
17	<u>47</u> (<u>13</u>)	43 (3)	62 (<u>14</u>)	<u>3</u>	<u>43</u>	<u>36</u>
6	<u>53</u> (7)	<u>36</u> (2)	<u>50</u> (9)	6	12	14
23	<u>56</u> (6)	52 (3)	<u>87</u> (1)	<u>45</u>	4	1
14	62 (2)	61 (1)	81 (5)	11	28	8
25	<u>64</u> (2)	53 (1)	70 (<u>12</u>)	<u>2</u>	29	10
24	<u>73</u> (7)	65 (6)	73 (5)	27	--	--

Sources: Fisher, Filby & Marliave, 1977 and Dishaw, 1977.

^aSelected numbers have been underlined to highlight them.

^bStandard deviations in parantheses.

Though prior achievement could explain some differences in time allocations (achievement data were unavailable for individual classrooms at time of the writing of this paper), it cannot be the sole source of differences in allocations and emphasis. The data reflect greater than a 3:1 ratio of differences in average daily time in total mathematics and larger differences in engagement rates and proportion of time devoted to specific topics. Classrooms 8 and 17 appear to be highly individualized in mathematics instruction with substantial differences in total time (33 and 47 minutes, respectively) and engagement rates (87% and 62%). Classroom 8 displays a more varied coverage of topics than classroom 17 (almost entirely multiplication and division). The data also suggests that uniform time allocations (as reflected in small standard deviations) do not necessarily lead to high or uniform engagement rates (compare, e.g., classrooms 14 and 25).

Variation in Teacher Behavior across Classrooms. The observed differences in time allocations and engagement rates are in part a function of teaching practices. Figure 1 presents schematically those teaching behaviors coded by FWL (see Marliave, Fisher, Filby & Dishaw, 1977 for details of observations). Table 4 presents data on teaching behavior for selected classrooms.

Figure 1 about here

Table 4 about here

Figure 1

Matrix of Combinations of Observation Categories

SETTING = Self-Paced

TEACHER BEHAVIOR			SOURCE				
	General Function	Specific Move	Teacher	Other Adult	Other Student	Curri- culum	
Proce- dural	Substantive	Monitoring				x	
			Academic Observation				x
			Academic Question				
			Academic Feedback				
	Feedback	Presentation	Explanation--Need				x
			Explanation--Planned				x
		Presentation	Structure, Direct				x
		Feedback	Engagement Feedback				x
Student FOCUS							

SETTING = Other-Paced

TEACHER BEHAVIOR			SOURCE				
		General Function	Specific Move	Teacher	Other Adult	Other Student	Curriculum
Proce- dural	Substantive	Monitoring	Academic Observation				x
			Academic Question				x
		Feedback	Academic Feedback				
		Presentation	Explanation--Need				x
			Explanation--Planned				x
	Presentation	Structure, Direct				x	
	Feedback	Engagement Feedback				x	
	Student Group						
FOCUS							

.x These combinations are not permissible within the observation system.

Source: Marliave, Fisher, Filby and Dishaw, 1977

Table 4. Percentages of Total Mathematics Time for Specific Teacher Behaviors during A-B and B-C Periods.^a

Other Paced												
Class	(Grade 5) Total Self Paced		Monitoring		Feedback		Presentation		Structure Direct		Engagement Feedback	
	A-B	B-C	A-B	B-C	A-B	B-C	A-B	B-C	A-B	B-C	A-B	B-C
3	65	69	*	4.2	11.6	7.1	6.0	6.2	7.5	5.8	7.5	2.7
4	74	49	5.3	11.6	*	10.5	15.1	11.0	*	5.1	*	*
8	98	87	*	*	*	*	*	*	*	3.4	*	*
18	88	94	*	*	6.9	2.3	*	*	*	*	*	*
17	88	90	*	*	*	*	4.5	2.2	*	3.5	*	*
6	86	81	*	*	3.0	*	4.2	2.6	*	*	*	*
23	52	56	5.8	4.7	8.0	19.2	10.0	6.5	2.2	2.0	*	*
14	50	52	8.1	8.8	16.8	5.2	5.6	8.8	6.3	3.6	2.1	*
25	96	91	*	*	*	*	3.0	7.2	*	*	*	*
24	96	86	*	*	3.6	*	*	11.7	*	*	*	*
MEAN	75	76	2.5	3.3	6.8	6.1	5.3	5.5	2.6	2.5	1.0	.6
RANGE	50-98	44-99	0-8.1	0-11.6	0-16.8	.5-19.2	0-15.7	.3-14.1	0-9.0	.2-5.8	0-7.5	0-2.7

Sources: Filby and Fisher, 1977; Filby, 1977.

^aAll behavior occurring less than 2% of the time are designated by an "*".

Apparently mathematics learning occurs predominantly in a self-paced setting regardless of the teacher's overall style of classroom instruction. Classrooms 14 and 23 seem to be the only consistent exceptions to this result. Class 4 reflects a large shift from self-paced to other-paced (teacher-paced) during the year.

The specific teacher behaviors reflected in the other-paced setting varied across classes and to some degree within classes for the two periods of observation. Persons well-grounded in appropriate pedagogical styles for specific curriculum topics might be able to explain the observed fluctuations.

Implications and Procedures for the Second Mathematics Study

I have tried to capture the lushness and diversity of the information about classrooms as gathered in the BTES project. FWL staff viewed themselves as being in clinical diagnostic roles and their practices reflected the desire to get beneath the surface of the manifest curriculum and global teacher style measures. The picture they came up with is complex and at times perplexing. It should be if the earlier argument about learning in a social context is at all accurate.

Recently, David Berliner (personal communication) did offer what he called a six step process for identifying good teaching. His steps, which he acknowledges greatly overlap Kounin's (1970) dimensions are:

- (1) Clear goals -- Can the researcher identify the goal of the lesson? If he can't, how can the student? Do students appear to understand what is expected of them? Are there directions and structuring kinds of events?
- (2) Appropriate time allocation -- Given the goal of the lesson, did the teacher allocate enough time to do the task?

- (3) Disjointedness -- Is the lesson disjointed? (student interruptions, etc.)
- (4) Engagement -- Is there evidence of engagement (attention and interest) in the classroom?
- (5) Monitoring -- Is the teacher providing lots of monitoring? (behavioral and academic)
- (6) Conviviality -- Amalgam of warmth, democratic style, manners

It seems to me that the IEA study should attempt to assess Berliner and Kounin's characteristics of classroom process within some reasonable constraints. These more detailed assessments should be limited to a small sample of classrooms and should include classroom observations on several occasions by, perhaps, multiple raters, teacher reports of time allocations for specific students in a small number of curriculum areas, ethnographic assessments of classroom processes, or some combination of the above.

My guess is that it is possible to do all three for a sample of classrooms where observations, teacher logs, and ethnographic studies are conducted within a constrained time period (less than a month) and focused on a couple of curriculum topics. The emphasis should be on those topics which are first introduced in the grade where the study is being conducted. Moreover, the intensive classroom process study should be timed to coincide with the teacher's normal coverage of one or more of the selected topics. Sample classes would potentially be studied at different times of the year and for only partially overlapping topics.

A study involving classroom observation, log keeping and ethnographic data collection would obviously require careful training (note the FWL descriptions of their training in the various reports) and careful coordination. The participating teachers and classrooms would have to be identified early, voluntary participation solicited, and the necessary training conducted.

These are all difficult tasks, but not insurmountable ones for a national center to carry out. For the sake of the improvement in study data and value for developing further expertise in national centers in these areas of intensive process measurement, I urge that the IMC and the participating countries give them strong consideration as a valuable adjunct to present plans.

FOOTNOTE

1. In actual practice, division and fraction subtests were not given at Time A. FWL used total multiplication score at time A as the "pretest" for all three time B achievement outcomes (multiplication, division, and fractions).

REFERENCES

- Berliner, D. C., Cahen, L. S., Filby, N. N., Fisher, C. W., Marliave, R. & Moore, J. Proposal for Phase III B of the Beginning Teacher Evaluation Study. Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, 1976.
- Bloom, B. S. Implications of the IEA studies for curriculum and instruction. School Review, May 1974, 413-435.
- Burstein, L. Analyzing multilevel educational data: The choice of analytical model rather than the units of analysis. Paper presented at the Winter Measurement and Methodology Conference, Center for the Study of Evaluation, University of California at Los Angeles, January 1978.
- Burstein, L. Three key topics in regression-based analysis of multilevel data from quasi-experiments and field studies. Paper presented at the Institute for Research on Teaching, Michigan State University, December, 1977.
- Burstein, L. The choice of unit of analysis in the investigation of school effects: The IEA in New Zealand. New Zealand Journal of Educational Studies, 1976, 11, 11-24.
- Burstein, L. & Smith, I. Choosing the appropriate unit for investigating school effects. The Australian Journal of Education, 1977, 21, 65-79.
- Carroll, J. B. The teaching of French as a foreign language in eight countries. International Studies in Evaluation, Vol. 5. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1975.
- Coleman, J. S. The effects of school on learning. Review of Educational Research, 1975, 45, 355-386.
- Comber, L. C. & Keeves, J. P. Science education in nineteen countries. International Studies in Evaluation, Vol. 1. Stockholm: Almqvist and Wiksell and New York: Wiley, 1973.
- Dishaw, M. Descriptions of allocated time to content areas for the B-C period. (Technical Note IV-2b) Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, August 1977.
- Dishaw, M. Descriptions of allocated time to content areas for the A-B period. (Technical Note IV-2a) Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, August, 1977.
- Filby, N. N. Descriptions of patterns of teaching behavior within and across classes during the A-B period. (Technical Note IV-3a) Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, August 1977.

- Filby, N. N. & Fisher, C. W. Description of patterns of teaching behavior within and across classes during the B-C period. (Technical Note IV-3b). Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, August 1977.
- Finn, J. D. & Kotlewski, L. D. Components of achievement in science. Paper presented at the Annual Meeting of the American Educational Research Association, New York, 1977.
- Fisher, C. W., Filby, N. N. & Marliave, R. S. Descriptions and distributions of ALT within and across classes during the B-C period. (Technical Note IV-1b). Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, October, 1977.
- Inkeles, A. A review of the IEA six-subjects survey. Proceedings of the National Academy of Education, 1976, 3.
- Kounin, J. S. Discipline and group management in classrooms. New York: Holt, Rinehart & Winston, 1970.
- Lav, L. Educational production functions. Paper presented at National Institute of Education conference on School Organization and Effects, San Diego, January 1978.
- Lewis, E. G. & Massad, C. E. The teaching of English as a foreign language in ten countries. International Studies in Evaluation, Vol. 4. Stockholm: Almqvist and Wiksell, and New York: John Wiley & Sons, 1975.
- Marliave, R. S., Fisher, C. W. & Dishaw, M. M. Academic learning time and student achievement in the A-B period (Technical Note V-1a) Beginning Teacher Evaluation Study, Far West Laboratory, San Francisco, August, 1977.
- Marliave, R. Fisher, C. W., Filby, N. N. & Dishaw, M. M. The development of instrumentation for a field study of teaching. (Technical Report I-5) Beginning Teacher Evaluation Study, Far West Laboratory, February 1977.
- Passow, A. H., Noah & Eckstein. The national case study: An empirical comparative study of twenty-one educational systems. International Studies in Evaluation, Vol. 6. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1976.
- Peaker, G. F. An empirical study of education in twenty-one countries: A technical report. Vol. 8. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1975.
- Postlethwaite, T. N., et. al. (Eds.) What do children know? A special issue on educational achievement. Comparative Education Review, 1974, vol. 18.
- Purves, A. C. Literature education in ten countries. International Studies in Evaluation, Vol. 2. Stockholm: Almqvist and Wiksell, and New York: John Wiley (Halsted Press), 1973.

Purves, A. C. & Levine, D. U. (Eds.) Educational policy and international assessment, Berkeley, CA.: McCutchan Publishers, Inc., 1975.

Schwille, J. S. Predictors of between-student differences in civic education education cognitive achievement. Chapter 6 in Torney, J.V., Oppenheim, A. N., & Farnen, R.F., Civic education in ten countries. International Studies in Evaluation, Vol. 6. Stockholm: Almqvist and Wiksell, and New York: John Wiley & Sons, 1975.

Thorndike, R. L. Reading comprehension education in fifteen countries. International Studies in Evaluation, Vol. 9. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1976.

Torney, J. V., Oppenheim, A. N. & Farnen, R. F. Civic education in ten countries. International Studies in Evaluation, Vol. 6. Stockholm: Almqvist & Wiksell, and New York: John Wiley & Sons, 1975.

Travers, K. J. The Second International Study of Mathematics: Purposes and design. Paper presented at the IEA meeting, National Institute of Educational Research, Tokyo, Japan, January 1978.

Walker, D. A. The IEA six-subject survey: An empirical study of education in twenty-one countries. International Studies in Evaluation, Vol. 9. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1976.